

# Security of practical private randomness generation

Stefano Pironio and Serge Massar

Laboratoire d'Information Quantique, Université Libre de Bruxelles, Brussels (Belgium)

March 4, 2011

## Abstract

Measurements on entangled quantum systems necessarily yield outcomes that are intrinsically unpredictable if they violate a Bell inequality. This property can be used to generate certified randomness in a device-independent way, i.e., without making detailed assumptions about the internal working of the quantum devices used to generate the random numbers. Furthermore these numbers are also private, i.e., they appear random not only to the user, but also to any adversary that might possess a perfect description of the devices. Since this process requires a small initial random seed to sample the behaviour of the quantum devices and to extract uniform randomness from the raw outputs of the devices, one usually speaks of device-independent randomness *expansion*.

The purpose of this paper is twofold. First, we point out that in most real, practical situations, where the concept of device-independence is used as a protection against unintentional flaws or failures of the quantum apparatuses, it is sufficient to show that the generated string is random with respect to an adversary that holds only classical-side information, i.e., proving randomness against quantum-side information is not necessary. Furthermore, the initial random seed does not need to be private with respect to the adversary, provided that it is generated in a way that is independent from the measured systems. The devices, though, will generate cryptographically-secure randomness that cannot be predicted by the adversary and thus one can, given access to free public randomness, talk about private randomness *generation*.

The theoretical tools to quantify the generated randomness according to these criteria were already introduced in *S. Pironio et al, Nature* **464**, 1021 (2010), but the final results were improperly formulated. The second aim of this paper is to correct this inaccurate formulation and therefore lay out a precise theoretical framework for practical device-independent randomness generation.

## 1 Introduction

Random numbers are essential for many applications such as computer simulations, statistical sampling, gambling, or video games. They are particularly important for classical and quantum cryptography, where the use of a flawed random number generator (RNG) can completely compromise the security. Many solutions have thus been proposed for the generation of random numbers (for recent on random number generation see, e.g., [1, 2, 3, 4, 5, 6, 7]), but none is entirely satisfactory. As quoted from Wikipedia [8], every random number generator (RNG) is subject to the following problems:

“It is very easy to misconstrue hardware or software devices which attempt to generate random numbers. Also, most ‘break’ silently, often producing decreasingly random numbers as they degrade. A physical example might be the rapidly decreasing radioactivity of the smoke detectors [...]. Failure modes in such devices are plentiful and are complicated, slow, and hard to detect.

Because many entropy sources are often quite fragile, and fail silently, statistical tests on their output should be performed continuously. Many, but not all, such devices include some such tests into the software that reads the device.

Just as with other components of a cryptosystem, a software random number generator should be designed to resist certain attacks. Defending against these attacks is difficult. [...]

[*On estimating entropy*]. There are mathematical techniques for estimating the entropy of a sequence of symbols. None are so reliable that their estimates can be fully relied upon; there are always assumptions which may be very difficult to confirm. These are useful for determining if there is enough entropy in a seed pool, for example, but they cannot, in general, distinguish between a true random source and a pseudo-random generator.

[*On performance test*]. Hardware random number generators should be constantly monitored for proper operation. [...] Unfortunately, with currently available (and foreseen) tests, passing such tests is not enough to be sure the output sequences are random. A carefully chosen design, verification that the manufactured device implements that design and continuous physical security to insure against tampering may all be needed in addition to testing for high value uses.”

Device-independent randomness generation aims to address these problems by exploiting the intrinsic unpredictability associated with the violation of Bell inequalities [9, 10, 11]. More precisely, consider a quantum system composed of two separated parts  $A$  and  $B$  which upon receiving respective inputs  $V^a$  and  $V^b$ , return respective outputs  $X^a$  and  $X^b$ . If after  $n$  successive uses of the devices, the *observed* data violates a Bell inequality, it is then possible to *certify* that the output string  $(X_1^a, X_1^b), \dots, (X_n^a, X_n^b)$  contains a certain amount of min-entropy, even when conditioned on the value of the inputs  $(V_1^a, V_1^b), \dots, (V_n^a, V_n^b)$ , and a randomness extractor can therefore be applied to the outputs to obtain almost-uniform random bits. Furthermore, this conclusion can be reached independently of any detailed assumptions about the inner working of the devices and is thus immune to most of the problems mentioned above.

That the violation of Bell inequalities is an indicator of quantum randomness had probably been recognized early on by many physicists, but was made explicit only recently in [10, 11, 12]. Not surprisingly, it was suggested shortly thereafter that Bell inequality violating systems could be exploited for randomness generation, and a scheme based on GHZ states was proposed in [13]. The possibility of device-independent randomness generation, however, was established only in [14], where a method to bound the min-entropy of the devices’ output as a function of the observed Bell violation was introduced. Furthermore, a proof-of-principle experimental demonstration was realised using two trapped ions.

The concept of device-independence (DI) is not restricted to randomness generation but includes adversarial applications such as quantum key distribution (QKD) [15, 16, 11, 17, 18] and coin tossing [19], and non-adversarial ones such as state estimation [20], entanglement witnesses [21], and self-testing of quantum computers [22]. In adversarial applications of device-independence it is often remarked that since the correctness of the protocol can be verified without making assumptions about the inner working of the devices, these could even have been prepared by the adversary itself. This has at least two implications as regards the theoretical analysis of device-independent randomness generation (and has also various implications for its experimental implementation, some of which will be briefly discussed later on).

First, if the adversary is allowed to prepare the quantum devices, nothing prevents him to entangle them with a quantum state that he keeps for himself in a quantum memory. It is then a priori possible that if he sees part of the devices’ output at some later stage, he could measure his quantum state in a way that would give him useful information about the

remainder of the output string. One thus needs to show that the output produced by the device also appear random with respect to the quantum-side information held by the adversary. The methods introduced in [14], however, have been shown so far to estimate randomness only against classical-side information, i.e., against adversaries who do not share entanglement with the quantum devices.

Second, if the adversary happens to have some prior knowledge of the inputs used to sample the devices, he could exploit it to program the devices in a way that would mimick the violation of a Bell inequality while at the same time giving him substantial information about the generated outputs. A random, private seed is thus necessary to select the inputs and start off the protocol. In addition, one also need some initial randomness to extract uniform random bits from the devices' outputs. One thus often speaks of device-independent randomness *expansion* (DIRE). A scheme achieving quadratic expansion was presented in [14], where it was also suggested to use more than one pair of devices to obtain greater (e.g., exponential) expansion.

In this paper, we do not investigate this extremal adversarial scenario where the quantum devices have been acquired from a malicious provider. We are instead interested in the more real-life and practical situation where the manufacturer of the device is assumed to be honest, but where the concept of device-independence is used to provide an accurate estimation of the amount of randomness generated independently of noise, limited control of the apparatuses, or unintentional flaws of the devices. We point out in Section 2 that in this context it is sufficient to prove security against classical-information. Furthermore, the initial seed used to sample the devices and perform the randomness extraction does not necessarily need to be private with respect to the adversary (it simply needs to be chosen in a way that is independent from the state of the devices). The output of the protocol, though, will represent a private random string. In this case one can thus talk about private randomness *generation*, given access to public randomness. (In the following, we will keep using the single terminology “device-independent randomness expansion” to refer to the two situations in which the initial randomness is considered to be private or is viewed as a free, public resource).

In Section 3, we then analyse the security of DIRE from this perspective. In particular, subsection 3.2. contains a detailed presentation of the model that we consider and of the assumptions on which it is based. Our main results are presented in subsection 3.3., where we show how to estimate the randomness produced in a Bell experiment if those assumptions are satisfied. Our analysis relies essentially on the tools introduced in [14], but importantly it fixes an issue that led to an improper formulation of the final results of [14]<sup>1</sup>. A very similar analysis has been presented in the independent work [23]. We briefly discuss how these results directly imply the security of various DIRE schemes in subsection 3.4.

Finally, we point out that a randomness-expansion scheme with superpolynomial expansion and proven to be secure against quantum side information was recently introduced in [24]. This protocol, however, requires an almost perfect violation of the CHSH inequality, while our results and those of [23] are generic and holds for arbitrary Bell inequalities and any amount of violation<sup>2</sup>.

## 2 Honest vs dishonest device suppliers and DIRE

The security of device-independent cryptographic protocols is based on a rather limited sets of assumptions, e.g., that the devices obey quantum theory, that separated devices can be

---

<sup>1</sup>Specifically, the problem lies with Eq. (3) and Eq. (A.9) of the Supplementary Information of [14] and with the final steps leading to these equations.

<sup>2</sup>Note that previous versions of these results (see [25] and [26]) claimed security against quantum side information, but both proofs were incorrect.

prevented to communicate with one other, that the users of the device have access to a private source of randomness, and so on. Provided that these basic assumptions are satisfied, the security follows independently of implementation details such as the precise quantum states and measurement operators used, or the dimension of the Hilbert space in which they are defined. It is often stressed that security could thus in particular be guaranteed if the devices had been provided or sabotaged by the adversary itself. This possibility is fascinating from a conceptual point of view and deserves to be investigated for its own sake. However, it has probably little (if no) practical relevance, as has already been pointed out (see e.g., [27]).

One reason is that while it is in principle possible to enforce the assumptions required for the security of a DI cryptography scheme based on malicious devices, in practice this may involve incredible technological and physical resources. For instance, how can we practically guarantee that the devices do not covertly leak out sensitive information to the adversary [14]? How can we guarantee that they do not contain sneaky transmitters? In principle communications through electromagnetic waves can be screened, but what about communications based on neutrinos or gravitational waves? When a “door” is opened to let a particle enter in a device, how can we efficiently prevent other particles to come out of the device?

More generally, any practical cryptographic implementation, classical, quantum, or device-independent, will include and make use of classical computing and communicating devices to process, store, and transmit data. These classical devices, which are probably easier to corrupt than their quantum counter-parts, cannot be guaranteed secure if they have been acquired from dishonest providers. One should therefore either acquire these classical devices from trusted suppliers or inspect them for malicious behaviour. But then why apply a different standard to the quantum devices?

The real problem, to which the concept of device-independence offers a potential solution, is that even if the quantum devices have been obtained from honest suppliers or thoroughly inspected, many things can still unintentionally go wrong. Indeed, in standard (i.e., device-dependent) quantum cryptography, conclusions about the randomness or the secrecy of the outputs crucially depends on the *physical* properties of the generation process, for example, on the fact that the outputs were produced by measuring the polarization of a single photon along well-defined directions. But then, how can one assess the level of security provided by a real-life implementation of a standard quantum cryptography protocol, which will inevitably differ in undetermined ways from the idealized, theoretical description [28]? Consider for instance that the reported attacks [29, 30, 31] on commercial QKD systems did not exploit any intentional, malicious flaws in the devices.

This problem is particularly acute in the case of (classical or quantum) RNG devices, as it is very difficult even for honest parties to construct reliable RNGs and monitor them for proper operation. The generation of randomness in a device-independent way solves many of the shortcomings of usual RNGs listed earlier, since it makes possible an accurate estimation of the amount of randomness generated independently of noise, imperfections, lack of knowledge, or limited control of the apparatuses.

The use of device-independence, even in a trusted provider situation, has the advantage over a full device-dependent approach that it requires only the verification of a limited number of precisely defined assumptions, on which the manufacturer of the device can focus. Furthermore, these assumptions can be much more easily enforced or verified with respect to the situation where the devices come from a dishonest provider, as one does not need to fight against devices that have been maliciously programmed<sup>3</sup> [14]. For instance, in the experiment reported in [14] no particular measures have been taken to screen-off one device from the other. However,

---

<sup>3</sup>Note in particular that it is highly unlikely that the attack reported in [32] would spontaneously occur in non-malicious devices.

the experiments involve two atoms that are confined in two independent vacuum chambers separated by about 1 meter. At this distance, direct interaction between the atoms is negligible and classical microwave and optical fields used to perform measurements on one atom have no influence on the other atom. Based on this superficial description of the setup, one can safely assume that the two quantum systems are independent and that no imperfections, failures, or implementation weaknesses would lead to direct interaction between the devices (though imperfections could lead to other potential problems that can be ruled out by the DI approach), and thus that the general formalism used to derive a bound on the randomness applies.

In the case of DIRE, assuming that the devices originate from a honest provider has not only experimental implications, but also theoretical ones. The first one is that, while the adversary may possess an arbitrarily accurate classical description of the internal working of the devices at any given moment of time, it is highly unlikely that he could possess any quantum system that is entangled with those inside the devices if he did not manufacture or tamper with them. This means that proving that the output are random with respect to *classical-side* information is sufficient.

The second implication is that the adversary cannot program the devices to exploit any prior knowledge about the initial randomness used to choose the inputs. The inputs must still be selected in a way that is independent from the internal functioning of the devices, but this condition can be satisfied without having recourse to cryptographically-secure random number generators. For instance, in the experiment reported in [14], the measurement settings were chosen by combining through a XOR function several public random number generators that use randomness derived from radioactive decay [33], atmospheric noise [34], and remote computer and network activity [35]. While a dishonest manufacturer aware of this procedure could have exploited it in the design of the set-up, it is highly unlikely that the state of the ions in the experiment of [14] was in any way correlated to the choice of measurement bases. If this condition is satisfied, it is justified, however, to conclude that the outputs of the devices do represent new, private random bits.

Remark that the two above implications are specific to DIRE but would not hold for most DI cryptographic protocols. This is due to the fact that DIRE is a single user protocol completely carried out in a single secure lab and which therefore does not allow for the possibility of interactive attacks by the adversary. In contrast, DIQKD, for instance, usually involves the sending of quantum information between Alice and Bob's devices. This quantum information can be intercepted by the adversary and entangled with his own quantum system. Furthermore any knowledge of the random numbers used in the protocol could be exploited by the adversary to improve the efficiency of this interaction. Even if the devices are completely trusted, it is therefore still the case that the security of QKD must be based on a proof that holds against quantum-side information and that the random numbers used in the protocol must be cryptographically secure.

In the following section we analyse DIRE from the perspective discussed above, and show in particular how to prove the security of a DIRE protocol against classical side-information.

### 3 DIRE against classical side-information

We start by recalling some definitions and results that will be used in the following. We refer to [14, 36, 37] for more details.

### 3.1 Preliminaries

**Random variables.** Let  $R$  be a random variable over the finite set  $\mathcal{R}$  and  $\Pr[R = r] = P_R(r)$  the probability that it takes the value  $r$ . (In the following, we use upper-case letters to denote random variable and lower-case letter to denote specific values taken by these variables). The closeness between two distributions  $P_R$  and  $Q_R$  can be quantified through the trace distance

$$d(P_R, Q_R) = \frac{1}{2} \sum_r |P_R(r) - Q_R(r)|. \quad (1)$$

For simplicity, we will write  $P(r)$  for the probabilities  $P_R(r)$ , when there is no risk of confusion. Let  $E$  be a random variable representing some classical side-information about the variable  $R$ , and let the correlations between  $R$  and  $E$  be described by a joint distribution  $\Pr[R = r, E = e] = P_{RE}(re)$ . We say that  $R$  is  $\delta$ -random with respect to  $E$  if it is  $\delta$ -close to a uniform distribution uncorrelated to  $E$ , that is if

$$d(P_{RE}, U_R \times Q_E) = \frac{1}{2} \sum_{r,e} |P_{RE}(re) - U_R(r) \times Q_E(e)| \leq \delta \quad (2)$$

for some distribution  $Q_E$ , where  $U_R(r) = 1/|\mathcal{R}|$  is the uniform probability distribution on  $\mathcal{R}$ .

**Min-entropy.** The randomness of  $R$  with respect to  $E$  can be quantified through the conditional min-entropy

$$H_{\min}(R|E)_P = -\log_2 \sum_{e \in \mathcal{E}} P_E(e) \max_{r \in \mathcal{R}} P_{R|E}(r|e). \quad (3)$$

The conditional min-entropy (3) is sometimes called the *average* conditional min-entropy to distinguish it from the *worst-case* conditional min-entropy defined by

$$\tilde{H}_{\min}(R|E)_P = -\log_2 \max_{r,e} P_{R|E}(r|e). \quad (4)$$

The worst-case min-entropy is a lower-bound on the average min-entropy:  $H_{\min}(R|E)_P \geq \tilde{H}_{\min}(R|E)_P$ . Note that when there is no side-information  $E$ , both entropies reduce to the usual definition  $H_{\min}(R)_P = -\log_2 \max_{r \in \mathcal{R}} P_R(r)$  for the classical min-entropy of a distribution  $P_R$ .

**Randomness extractors.** Given a  $n$ -bit string  $R$  with a certain conditional min-entropy  $k$  one can extract from it, using a randomness extractor and a small uniform seed  $S$ , a new  $m$ -bit random string that is almost uniformly random. More formally, a function  $\text{Ext} : \{0, 1\}^n \times \{0, 1\}^d \rightarrow \{0, 1\}^m$  is a  $(m, k, \delta)$ -strong extractor with uniform seed if for all distributions  $P_{RE}$  with  $H_{\min}(R|E)_P \geq k$ , and for a uniform seed  $S \in \{0, 1\}^d$ , we have<sup>4</sup>

$$d(P_{\text{Ext}(R,S)SE}, U_m \times P_S \times P_E) \leq \delta, \quad (5)$$

where  $U_m$  is the uniform distribution on  $\{0, 1\}^m$ . There exist different construction for randomness extractors, characterized by different relations between the parameters  $n, m, d, k, \delta$ . In particular, for any  $k$  and  $\delta$ , there exist extractors with output length  $m = k - 4 \log 1/\delta - O(1)$  and seed length  $d = O(\log^2(n/\delta) \log m)$  [37].

**Randomness and Bell experiments.** In [14], it was shown that there exists a fundamental, quantitative relation between the violation of Bell inequalities and the randomness produced in Bell experiments. We consider here for simplicity Bell experiments performed on two distinct systems  $A$  and  $B$ , although our results generalize to more parties. We denote  $V = (V^a, V^b)$  the

---

<sup>4</sup>Note that the definition of (classical) extractors does not usually involve side-information, but the definition given here and the conventional one can be shown to be essentially equivalent [38].



measurement choices for systems  $A$  and  $B$  and assume that they each take values in a finite set  $\mathcal{V}$ . We denote the measurement outputs  $X = (X^a, X^b)$  and assume that they each take values in the finite set  $\mathcal{X}$ . To any given input  $V^a = v^a$ , we can associate a set of measurement operators  $\{M_A(x^a|v^a)\}_{x^a \in \mathcal{X}}$  such that  $\sum_{x^a} M_A^\dagger(x^a|v^a) M_A(x^a|v^a) = I_A$ , where  $I_A$  is the identity operator on the Hilbert space  $\mathcal{H}_A$  of system  $A$ . Similarly a set of measurement operators  $M_B(x^b|v^b)$  can be associated to any given input  $V^b = v^b$ . The probability to obtain the pair of outputs  $x = (x^a, x^b)$  given the pair of inputs  $v = (v^a, v^b)$  when measuring a joint state  $\rho_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$  can then be written

$$P(x|v) = \text{tr} \left[ M_A(x^a|v^a) \otimes M_B(x^b|v^b) \rho_{AB} M_A^\dagger(x^a|v^a) \otimes M_B^\dagger(x^b|v^b) \right]. \quad (6)$$

A Bell expression  $I$  is defined by a series of coefficients  $c_{vx}$ , which associate to a conditional probability distribution  $P = \{P(x|v)\}$  the *Bell expectation*  $I[P] = \sum_{vx} c_{vx} P(x|v)$ . We denote by  $I_q$  the maximal quantum Bell expectation, i.e.,  $I_q = \max_P I[P]$ , where the maximum is taken over all distributions of the form (6).

In [14] (see also [39]), it is shown that there exists a fundamental relation between the randomness of the distribution  $P$  and the Bell expectation  $I[P]$ . More precisely, it is shown how using the semidefinite programming hierarchy introduced in [40, 41], one can compute for each  $v$  a bound of the form

$$\max_x P(x|v) \leq g(I[P]), \quad (7)$$

which is valid for any state  $\rho_{AB}$  and measurement operators  $M_A(x^a|v^a)$ ,  $M_B(x^b|v^b)$  such that (6) holds. Here  $g$  is a function that is concave (if not, we take its concave hull) and monotonically decreasing, taking values between 1 and  $1/|\mathcal{X}|^2$ . In particular, it is thus also logarithmically concave. The above bound can be rewritten as  $H_{\min}(X|V=v)_P \geq f(I[P])$  where  $H_{\min}(X|V=v)_P = -\log_2 \max_x P(x|v)$  is the min-entropy of  $X$  for given  $v$ , and  $f(I[P]) = -\log_2 g(I[P])$ . From now on we refer to  $g$  (or  $f = -\log_2 g$ ) as a randomness bound associated to  $I$ .

### 3.2 Modelling of the devices and basic assumptions

We consider a single pair of Bell violating devices  $A$  and  $B$  (though the results below can be directly generalized to a multipartite setting), in which, the user Alice can respectively introduce inputs  $V = (V^a, V^b)$  (the “measurement settings”) and obtains output  $X = (X^a, X^b)$  (the “measurement outcomes”). The quantum apparatuses are used  $n$  times in succession for varying choices of the inputs. In full generality, the behaviour of the devices can be characterized by

- an initial state  $\rho_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$ ;
- a set  $\mathcal{M}_{AB} = \{M_{AB}(x|v)\}$  of measurement operators on  $\mathcal{H}_A \otimes \mathcal{H}_B$ , which have the product form

$$M_{AB}(x|v) = M_A(x^a|v^a) \otimes M_B(x^b|v^b), \quad (8)$$

and which define the measurements applied on the state of the devices for given input  $v = (v^a, v^b)$ ;

- a joint unitary operation  $U \in \mathcal{H}_A \otimes \mathcal{H}_B$ , which is applied on the post-measurement state of the devices after each measurement and which represents the possibility for the devices to communicate between successive measurements (e.g., to establish new entanglement).

Note that to simplify the notation, we did not explicitly introduce a dependence of  $M_{AB}(x|v)$  or  $U$  on the measurement round  $i$  or on the inputs and outputs obtained in previous steps, i.e.,

$M_{AB}(x|v)$  and  $U$  are identical at each use of the devices. The above formulation is nevertheless completely general and can account for the possibility that the behaviour of the devices varies from one round to another and makes use of an internal memory. Indeed, the measurement operators  $M_{AB}(x|v)$  and the operation  $U$  can encode the value of the inputs  $v$  and the output  $x$  obtained in a given run in the post measurement state of the devices and “read” back this information in the next step to perform an operation conditional on the previous history. The only restrictive hypothesis that we make is that the measurement operators have the product form (8). Physically, this means that the systems  $A$  and  $B$  do not communicate with each other during the measurement itself.

We assume that the behaviour of the devices, characterized by the initial state  $\rho_{AB}$ , the set of measurement operators  $\mathcal{M}_{AB}$ , and the joint operation  $U$ , is perfectly known to the adversary. Note that the behaviour of the devices might depend on some external random parameters known or controlled by the adversary. For instance, the quality of the components used to produce the devices might vary in a way known to the adversary or he might control some parameters (such as temperature or changes in the voltage of the power supply) that can influence the output of the devices. This can be taken into account by assuming that the devices and the adversary’s information are in a joint state

$$\rho_{ABE} = \sum_e P(e) \rho_{AB}^e \otimes |e\rangle\langle e|, \quad (9)$$

where  $\rho_{AB} = \sum_e P(e) \rho_{AB}^e$  and  $e$  represents the knowledge that the adversary has on the state of the devices. We refer in the following to  $(\rho_{ABE}, \mathcal{M}_{AB}, U_{AB})$  as the *device behaviour*. Our assumption of classical side-information lies in the fact that the devices and the adversary are only classically correlated. In general, i.e., in the case of quantum side-information, the state  $\rho_{ABE}$  could be completely arbitrary.

As we said, the devices will be used  $n$  times in succession. Let  $\mathbf{V} = (V_1, \dots, V_n) = (V_1^a, V_1^b, \dots, V_n^a, V_n^b)$  denote the sequence of inputs employed in  $n$  such successive uses and let  $P(\mathbf{v})$  denote the probability of a particular sequence  $\mathbf{V} = \mathbf{v}$ . We assume that the choice of inputs is independent of the device behaviour, i.e., that the inputs  $\mathbf{V}$ , the pair of devices  $AB$ , and the adversary’s information  $E$  can initially be characterized by the *cqc*-state

$$\rho_{\mathbf{V}} \otimes \rho_{ABE} = \sum_{\mathbf{v}, e} P(\mathbf{v}) P(e) |\mathbf{v}\rangle\langle \mathbf{v}| \otimes \rho_{AB}^e \otimes |e\rangle\langle e|. \quad (10)$$

After the  $n$  uses of the devices, one obtains a sequence  $\mathbf{X} = (X_1^a, X_1^b, \dots, X_n^a, X_n^b)$  of output pairs. The resulting situation, and the correlations between the inputs  $\mathbf{V}$ , outputs  $\mathbf{X}$ , and the adversary’s information  $E$ , can then be characterized by the joint distribution

$$P(\mathbf{vxe}) = P(\mathbf{v}) P(e) P(\mathbf{x}|\mathbf{v}, e), \quad (11)$$

where

$$P(\mathbf{x}|\mathbf{v}, e) = \text{tr} \left[ \prod_{i=1}^n \left( U_{AB} M_A(x_i^a|v_i^a) \otimes M_B(x_i^b|v_i^b) \right) \rho_{AB}^e \prod_{i=1}^n \left( M_A^\dagger(x_i^a|v_i^a) \otimes M_B^\dagger(x_i^b|v_i^b) U_{AB}^\dagger \right) \right] \quad (12)$$

represents the response of the devices to given inputs  $\mathbf{v}$  for a given value of the adversary’s information  $e$ .

In the following, we show how the level of Bell violation which is observed after  $n$  repetitions of the experiment implies a bound on the min-entropy of the output string  $\mathbf{X}$  conditioned on the input string  $\mathbf{V}$  and the adversary’s information  $E$ . This bound depends only on the product



assumption (8) characterizing the two devices, on the independence assumption (10) between the choice of inputs and the state of the devices, and implicitly on the condition (9) that the adversary's side-information is classical. Apart from these three assumptions, our results do not depend on any specific details of the device behaviour  $(\rho_{ABE}, \mathcal{M}_{AB}, U_{AB})$ .

### 3.3 Bounding the min-entropy

Suppose that the sequence of inputs  $\mathbf{V} = (V_1^a, V_1^b, \dots, V_n^a, V_n^b)$  is generated by choosing each pair of inputs  $(V_i^a, V_i^b)$  independently with probability  $\Pr[V_i^a = v, V_i^b = w] = p_{vw}$ , with  $q = \min_{v,w} p_{vw} > 0$ . Let  $I$  be a Bell expression  $I$  adapted to the input and output alphabet of the quantum devices. We then introduce the following Bell estimator

$$\bar{I} = \frac{1}{n} \sum_{i=1}^n I_i, \quad (13)$$

where

$$I_i = \sum_{xyvw} c_{xyvw} \frac{\chi(X_i^a = x, X_i^b = y, V_i^a = v, V_i^b = w)}{p_{vw}}. \quad (14)$$

Here,  $\chi(e)$  is the indicator function for the event  $e$ , that is,  $\chi(e) = 1$  if the event  $e$  is observed,  $\chi(e) = 0$  otherwise. The series of coefficients  $c_{xyvw}$  in (14) define the Bell expression  $I$ . We assume that they satisfy  $c = \max_{x,y,v,w} c_{xyvw} < \infty$ .

Let  $\{J_m : 0 \leq m \leq m_{\max}\}$  be a series of Bell violation thresholds, with  $J_0$  corresponding to the local bound of the Bell expression and  $J_{\max} = I_q$  to the maximum violation allowed by quantum theory. We are going to put a bound on the min-entropy of the string  $\mathbf{X}$  conditioned on the fact that the observed Bell average value  $\bar{I}$  is comprised within some interval<sup>5</sup>  $J_m \leq \bar{I} < J_{m+1}$ . We denote  $P(m)$  the probability that the experiment returns a Bell average value comprised between  $J_m \leq \bar{I} < J_{m+1}$  and  $H_{\min}(\mathbf{X}|\mathbf{V}E, m)_P$  the min-entropy of  $\mathbf{X}$  conditioned on  $\mathbf{V}$  and  $E$  given that a specific value  $m$  has been obtained. The case  $m = 0$  corresponds to the situation where no substantial Bell violations is observed and no randomness is produced.

**Theorem 1.** *Suppose that the sequence of inputs  $\mathbf{V} = (V_1^a, V_1^b, \dots, V_n^a, V_n^b)$  is generated by choosing each pair of inputs  $(V_i^a, V_i^b)$  independently with probability  $\Pr[V_i^a = v, V_i^b = w] = p_{vw}$ , with  $q = \min_{v,w} p_{vw} > 0$ . Let  $\epsilon, \epsilon' > 0$  be two arbitrary parameters. Then for any device behaviour  $(\rho_{ABE}, \mathcal{M}_{AB}, U)$ , the resulting distribution  $P = \{P(\mathbf{v}|\mathbf{x}e)\}$  characterizing  $n$  successive use of the devices is  $\epsilon$ -close to a distribution  $Q$  such that*

1. *either  $Q(m) \leq \epsilon'$ ,*
2. *or  $H_{\min}(\mathbf{X}|\mathbf{V}E, m)_Q \geq nf(J_m - \mu) - \log_2 \frac{1}{\epsilon'}$ ,*

where  $f$  is a randomness bound associated to the Bell expression  $I$  and

$$\mu = \left( \frac{c}{q} + I_q \right) \sqrt{\frac{2}{n} \ln \frac{1}{\epsilon}}. \quad (15)$$

This results tell us that the classical distribution  $P$  characterizing the outputs  $\mathbf{X}$  of the devices and their correlations with the inputs  $\mathbf{V}$  and the adversary's information  $E$  is essentially indistinguishable from a distribution  $Q$  such that if the observed violation lies within the interval  $J_m \leq \bar{I} < J_{m+1}$  with non-negligible probability, then we have the guarantee that the outputs contain a certain amount of entropy, roughly given by  $nf(J_m)$  up to epsilonic corrections (remark

---

<sup>5</sup>This is the novel ingredient that fixes the issue in [14].

the term  $-\log_2 1/\epsilon'$  in the bound on the min-entropy  $H_{\min}(\mathbf{X}|\mathbf{V}E, m)_Q$  which was missing in [14]). Note that the fact that the trace distance cannot increase under classical-processing operations guarantees that any claim about the string  $\mathbf{X}$  (or any subsequent use thereof) which is based on the properties of the distribution  $Q$  will also hold for the distribution  $P$  up to a correction  $\epsilon$  (see subsection 3.4 for more details).

**Proof of Theorem 1.** In the following, we write  $\mathbf{v}_i = (v_1^a, v_1^b, \dots, v_i^a, v_i^b)$  for the collection of input pairs up to round  $i$ , and similarly for  $\mathbf{x}_i$ . We denote  $\mathbb{E}(I_i|\mathbf{x}_{i-1}, \mathbf{v}_{i-1}, e)$  the expectation of the random variable  $I_i$  defined in (14) conditioned on  $(\mathbf{x}_{i-1}, \mathbf{v}_{i-1}, e)$ , where the expectation is taken with respect to the probability distribution  $P$ . The following Lemma puts a bound on the probabilities  $P(\mathbf{x}|\mathbf{v}, e)$ .

**Lemma 1.** *Let  $G_\mu = \{(\mathbf{x}, \mathbf{v}, e) \mid \frac{1}{n} \sum_{i=1}^n \mathbb{E}(I_i|\mathbf{x}_{i-1}, \mathbf{v}_{i-1}, e) \geq \bar{I}(\mathbf{x}, \mathbf{v}) - \mu\}$ , where  $\mu \in \mathbb{R}$  is some real parameter. Then for any  $(\mathbf{x}, \mathbf{v}, e) \in G_\mu$ ,*

$$P(\mathbf{x}|\mathbf{v}, e) \leq g^n(\bar{I}(\mathbf{x}, \mathbf{v}) - \mu). \quad (16)$$

*Proof.* Using successively Bayes's rule and (12), we can write

$$P(\mathbf{x}|\mathbf{v}, e) = \prod_{i=1}^n P(x_i|v_i, \mathbf{x}_{i-1}, \mathbf{v}, e) = \prod_{i=1}^n P(x_i|v_i, \mathbf{x}_{i-1}, \mathbf{v}_{i-1}, e). \quad (17)$$

The second equality simply expresses the fact that the outputs at round  $i$  are determined only by the inputs at round  $i$  and by the past inputs and outputs, but not by future inputs. Note furthermore that we can write

$$\begin{aligned} P(x_i|v_i, \mathbf{x}_{i-1}, \mathbf{v}_{i-1}, e) &= P(x_i^a, x_i^b|v_i^a, v_i^b, \mathbf{x}_{i-1}, \mathbf{v}_{i-1}, e) \\ &= \text{tr}[M_A(x_i^a|v_i^b) \otimes M_B(x_i^b|v_i^a) \rho_{AB}^{e, \mathbf{x}_{i-1}, \mathbf{v}_{i-1}} M_A^\dagger(x_i^a|v_i^b) \otimes M_B^\dagger(x_i^b|v_i^a)], \end{aligned} \quad (18)$$

where  $\rho_{AB}^{e, \mathbf{x}_{i-1}, \mathbf{v}_{i-1}}$  denotes the state of the devices conditioned on previous inputs and outputs. Applying the randomness bound (7) to the probability distribution  $P_{\mathbf{x}_{i-1}, \mathbf{v}_{i-1}, e} = \{P(x_i|v_i, \mathbf{x}_{i-1}, \mathbf{v}_{i-1}, e)\}$  implies that  $P(x_i|v_i, \mathbf{x}_{i-1}, \mathbf{v}_{i-1}, e) \leq g(I[P_{\mathbf{x}_{i-1}, \mathbf{v}_{i-1}, e}])$ . Using the fact that  $P(v_i^a = v, v_i^b = w|\mathbf{x}_{i-1}, \mathbf{v}_{i-1}, e) = p_{vw}$  which follows from (10) and the fact that each pair of inputs  $(V_i^a, V_i^b)$  is generated independently with probability  $\Pr[V_i^a = v, V_i^b = w] = p_{vw}$ , it is easily verified that  $I[P_{\mathbf{x}_{i-1}, \mathbf{v}_{i-1}, e}] = \sum_{xyvw} c_{xyvw} P(xy|vw, \mathbf{x}_{i-1}, \mathbf{v}_{i-1}, e) = \mathbb{E}(I_i|\mathbf{x}_{i-1}, \mathbf{v}_{i-1}, e)$ . We therefore have

$$P(\mathbf{x}|\mathbf{v}, e) \leq \prod_{i=1}^n g(\mathbb{E}(I_i|\mathbf{x}_{i-1}, \mathbf{v}_{i-1}, e)) \leq g^n\left(\frac{1}{n} \sum_{i=1}^n \mathbb{E}(I_i|\mathbf{x}_{i-1}, \mathbf{v}_{i-1}, e)\right) \quad (19)$$

where we used that  $g$  is logarithmically concave in the second inequality. Using the definition of  $G_\mu$  and the fact that  $g$  is monotonically decreasing, we get (16).  $\square$

**Lemma 2.** *For any  $\epsilon > 0$ , let*

$$\mu = \left(\frac{c}{q} + I_q\right) \sqrt{\frac{2}{n} \ln \frac{1}{\epsilon}}. \quad (20)$$

*Then*

$$\Pr[G_\mu] = \sum_{(\mathbf{x}, \mathbf{v}, e) \in G_\mu} P(\mathbf{x}, \mathbf{v}, e) \geq 1 - \epsilon. \quad (21)$$

*Proof.* Consider the list of random variables  $Z_0, \dots, Z_n$ , where  $Z_0 = 0$  and

$$Z_k = \sum_{i=1}^k (I_i - \mathbb{E}(I_i | W_{k-1})) \quad (22)$$

for  $k \geq 1$ , where  $W_{k-1} = (\mathbf{X}_{k-1}, \mathbf{V}_{k-1}, E)$  and  $W_0 = E$ . Since  $|I_k| \leq c/q$  with  $c < \infty$  and  $q > 0$ , we have that  $|Z_k| \leq 2kc/q < \infty$  is bounded for all  $k$ . Moreover, the differences  $|Z_{k+1} - Z_k|$  are bounded by  $|Z_{k+1} - Z_k| = |I_{k+1} - \mathbb{E}(I_{k+1} | W_k)| \leq |I_{k+1}| + |\mathbb{E}(I_{k+1} | W_k)| \leq c/q + I_q$ , where we used (10) and the fact that each pair of inputs  $(V_i^a, V_i^b)$  is generated independently with probability  $\Pr[V_i^a = v, V_i^b = w] = p_{vw}$ . Finally, it is easily verified that  $\mathbb{E}(Z_{k+1} | W_k) = Z_k$  for all  $0 \leq k \leq n-1$ . The variables  $Z_0, \dots, Z_n$  thus form a martingale with respect to (the filtration induced by)  $W_0, \dots, W_{n-1}$ . We can therefore apply Azuma-Hoeffding inequality [42], which yields

$$\Pr[Z_n - Z_0 \geq n\mu] = \Pr\left[\frac{1}{n} \sum_{i=1}^n \mathbb{E}(I_i | W_{i-1}) \leq \bar{I} - \mu\right] \leq \exp\left(\frac{-n\mu^2}{2(c/q + I_q)^2}\right) = \epsilon, \quad (23)$$

which gives the desired claim given the definition of  $G_\mu$ .  $\square$

So far, we have (implicitly) considered the random variable sequence  $\mathbf{X}$  as taking value in the output space  $\mathcal{X}^n = \mathcal{X} \times \dots \times \mathcal{X}$ . We now formally extend its range and view it as an element of  $\mathcal{X}^n \cup \perp$  (with  $P(\mathbf{x} | \mathbf{v}, e) = 0$  if  $\mathbf{x} = \perp$ ). We can interpret  $\perp$  as an “abort-output” produced by the devices implying that no violation has been obtained (i.e.  $m = 0$  if  $\mathbf{x} = \perp$ ).

**Lemma 3.** *There exists a probability distribution  $Q = \{Q(\mathbf{x}, \mathbf{v}, e)\}$  that is  $\epsilon$ -close to  $P$  satisfying*

$$Q(\mathbf{x} | \mathbf{v}, e) \leq g^n(\bar{I}(\mathbf{x}, \mathbf{v}) - \mu). \quad (24)$$

for all  $(\mathbf{x}, \mathbf{v}, e)$  such that  $\mathbf{x} \neq \perp$ , with  $\mu$  given by (20).

*Proof.* Define  $Q$  as  $Q(\mathbf{x}, \mathbf{v}, e) = P(\mathbf{v})P(e)Q(\mathbf{x} | \mathbf{v}, e)$ , where  $Q(\mathbf{x} | \mathbf{v}, e) = P(\mathbf{x} | \mathbf{v}, e)$  if  $(\mathbf{x}, \mathbf{v}, e) \in G_\mu$ ,  $Q(\mathbf{x} | \mathbf{v}, e) = 0$  if  $\mathbf{x} \neq \perp$  and  $(\mathbf{x}, \mathbf{v}, e) \notin G_\mu$ , and  $Q(\perp | \mathbf{v}, e) = 1 - \sum_{\mathbf{x} \neq \perp} P(\mathbf{x} | \mathbf{v}, e)$ . By Lemma 1, the distribution  $Q$  satisfies (24) for all  $(\mathbf{x}, \mathbf{v}, e)$  such that  $\mathbf{x} \neq \perp$ . Application of Lemma 2 gives  $d(P, Q) = \frac{1}{2} \sum_{\mathbf{x}, \mathbf{v}, e} |P(\mathbf{x}, \mathbf{v}, e) - Q(\mathbf{x}, \mathbf{v}, e)| = \frac{1}{2} \sum_{\mathbf{v}, e} P(\mathbf{v}, e) \sum_{\mathbf{x}} |P(\mathbf{x} | \mathbf{v}, e) - Q(\mathbf{x} | \mathbf{v}, e)| = \frac{1}{2} (\sum_{\mathbf{x}, \mathbf{v}, e \notin G_\mu} P(\mathbf{x}, \mathbf{v}, e) + 1 - \sum_{\mathbf{x}, \mathbf{v}, e \in G_\mu} P(\mathbf{x}, \mathbf{v}, e)) \leq \epsilon$ .  $\square$

Let  $Q(m)$  be the probability (according to the distribution  $Q$ ) that  $J_m \leq \bar{I} < J_{m+1}$ . Let  $Q(\mathbf{x}, \mathbf{v}, e | m)$  denote the distribution of  $\mathbf{X}, \mathbf{V}, E$  conditioned on a particular value of  $m$  and let

$$H_{\min}(\mathbf{X} | \mathbf{V}, E, m)_Q = -\log_2 \sum_{\mathbf{v}, e} Q(\mathbf{v}, e | m) \max_{\mathbf{x}} Q(\mathbf{x} | \mathbf{v}, e, m) \quad (25)$$

be the min-entropy of the raw string  $\mathbf{X}$  conditioned on  $(\mathbf{V}, E)$  for a given  $m$ . Let  $K_m = \{\mathbf{x} | \mathbf{x} \neq \perp \text{ and } J_m \leq \bar{I}(\mathbf{x}, \mathbf{v}) < J_{m+1}\}$ . By Lemma 3 and the fact that the  $g$  is monotonically decreasing, we have

$$\max_{\mathbf{x}} Q(\mathbf{x} | \mathbf{v}, e, m) = \frac{1}{Q(m | \mathbf{v}, e)} \max_{\mathbf{x} \in K_m} Q(\mathbf{x} | \mathbf{v}, e) \quad (26)$$

$$\leq \frac{g^n(J_m - \mu)}{Q(m | \mathbf{v}, e)}. \quad (27)$$

Inserting this back in (25) gives

$$H_{\min}(\mathbf{X}|\mathbf{V}E, m)_Q \geq -\log_2 \sum_{\mathbf{v}, e} \frac{Q(\mathbf{v}, e|m)}{Q(m|\mathbf{v}, e)} g^n(J_m - \mu) \quad (28)$$

$$= -\log_2 \sum_{\mathbf{v}, e} \frac{Q(\mathbf{v}, e)}{Q(m)} g^n(J_m - \mu) \quad (29)$$

$$= nf(J_m - \mu) - \log_2 \frac{1}{Q(m)}, \quad (30)$$

where we remind that  $f = -\log_2 g$ . This immediately implies Theorem 1.

### 3.4 Application to DIRE protocols

Theorem 1 can directly be applied to prove the security of various DIRE protocols. Formally, a randomness expansion protocol is a protocol that, starting from a  $d$ -bit uniform random seed  $S$ , generates a  $m$ -bit string  $R$  that is close to uniformly random and uncorrelated from any potential adversary. The length  $m$  of the output string is variable and determined during the run of the protocol. The protocol may also abort, in which case we set  $m = 0$  and  $R = \emptyset$ . We can assume that  $m$  is made public at the end of the protocol.

The protocol will involve the use of Bell-violating devices and some classical processing on the outputs of the devices. For example a straightforward protocol directly based on the simple Bell experiment described so far is described below. But one could also consider more complicated protocols involving multiple pairs of Bell-violating devices, where this simple primitive is repeated or concatenated, see [25, 26].

1. **Input generation:** Alice generates a sequence of input pairs  $\mathbf{V} = (V_1^a, V_1^b, \dots, V_n^a, V_n^b)$  according to the (non-uniform) distribution specified in the statement of Theorem 1. This can be achieved starting from a uniform random seed  $S_{\text{inp}}$  with a small error  $\epsilon_{\text{inp}}$  and small entropy loss (see [45, 46] and the Appendix).
2. **Use of the devices:** She introduces inputs  $V_i^a$  and  $V_i^b$  in the two devices and obtains outputs  $X_i^a$  and  $X_i^b$ . This step is repeated  $n$  times, resulting in the sequence of output pairs  $\mathbf{X} = (X_1^a, X_1^b, \dots, X_n^a, X_n^b)$ .
3. **Estimation of the Bell expression:** Alice computes the average Bell expression (13) and determines the value of  $m$  such that  $J_m \leq \bar{I} < J_{m+1}$ . If  $m = 0$ , she aborts.
4. **Randomness extraction:** Using a random seed  $S_{\text{ext}}$ , Alice applies a  $(m, k_m, \epsilon_{\text{ext}})$ -randomness extractor to the raw string  $\mathbf{X}$  with  $k_m = nf(J_m - \mu) - \log_2 m_{\text{max}} - \log_2 \frac{1}{\epsilon'}$  and obtains a string  $R = \text{Ext}(\mathbf{X}, S_{\text{ext}})$ , which represents the output of the protocol. We can assume that  $m$ ,  $\mathbf{V}$ , and  $S_{\text{ext}}$  are made public.

In the above description, we have of course implicitly assumed that the thresholds  $J_m$ , the parameter  $\epsilon$  (which determine  $\mu$ ),  $\epsilon'$ , and  $\epsilon_{\text{ext}}$  are chosen in such a way that they define a proper  $(m, k_m, \epsilon_{\text{ext}})$ -randomness extractor for all values of  $m = 1, \dots, m_{\text{max}}$ .

Let  $F = (\mathbf{V}, S_{\text{ext}}, E)$  denote the final side-information of the adversary. Following the definition of security in the context of quantum key distribution outlined in [43, 44], we say that a protocol such as the one just presented is secure if, for any device behaviour and any  $m$ , the output  $R$  is uniformly random and independent from  $F$ . This means that the distribution

$P_{RFM}^{\text{perf}}$  characterizing the output  $R$ , the side-information  $F$  and the final length  $M$  of a perfectly secure protocol has the form

$$P_{RFM}^{\text{perf}}(rfm) = P_M(m) \times P_{RF|M}(rf|m) \quad \text{with} \quad P_{RF|M}(rf|m) = U_m(r) \times P_{f|M}(f|m) \quad (31)$$

where  $U_m$  is the uniform distribution on  $\{0, 1\}^m$ . A real DIRE protocol is said to be  $\epsilon_{\text{sec}}$ -secure if it is  $\epsilon_{\text{sec}}$ -indistinguishable from a secure protocol, that is, if for any device behaviour, the joint distribution  $P_{RFM}$  satisfies

$$d(P_{RFM}, P_{RFM}^{\text{perfect}}) \leq \epsilon_{\text{sec}} \quad (32)$$

for some distribution  $P_{RFM}^{\text{perfect}}$  of the form (31). In particular, a DIRE protocol is  $\epsilon_{\text{sec}}$ -secure if, for any device behaviour, it outputs  $m$ -bit strings that are  $\delta_m$ -random with respect to  $E$  with

$$\sum_{m=1}^{m_{\max}} p_m \delta_m \leq \epsilon_{\text{sec}}, \quad (33)$$

where  $m_{\max}$  denotes the maximal output length.

To show that the protocol defined above is secure according to this definition, suppose that at the end of Step 2, after the  $n$  uses of the devices, the correlations between outputs  $\mathbf{X}$ , inputs  $\mathbf{V}$ , and the adversary's prior information  $E$  are characterized by the probability distribution  $Q_{\mathbf{X}\mathbf{V}E}$  defined in the statement of Theorem 1. Then it is easy to show that the distribution  $Q_{RFM} = Q_{G(\mathbf{X}, \mathbf{V}, S_{\text{ext}})FM}$  characterizing the final output of the protocol (where  $G$  is the classical processing describing the steps performed after the  $n$  uses of the devices) is  $(\epsilon' + \epsilon_{\text{ext}})$ -close to a perfectly secure distribution  $\tilde{Q}_{RFM}$ . Indeed, let  $M_{<}$  be the values of  $m$  such that  $Q(m) \leq \epsilon'/m_{\max}$  and  $M_{>}$  those for which  $Q(m) \geq \epsilon'/m_{\max}$ . For all  $m \in M_{>}$ , the min-entropy  $H_{\min}(\mathbf{X}|\mathbf{V}E, m)_Q$  can thus be bounded by

$$H_{\min}(\mathbf{X}|\mathbf{V}E, m)_Q \geq nf(J_m - \mu) - \log_2 m_{\max} - \log_2 \frac{1}{\epsilon'}. \quad (34)$$

Applying a  $(m, k_m, \epsilon_{\text{ext}})$ -randomness extractor to the string  $\mathbf{X}$  with  $k_m$  given by the right-hand side of (34) therefore yields a string that is  $\delta_m$ -close to a random string, with  $\delta_m \leq \epsilon_{\text{ext}}$  for  $m \in M_{>}$  and  $\delta_m \leq 1$  for  $m \in M_{<}$ . On average, we thus have

$$\sum_m Q(m) \delta_m \leq \sum_{m \in M_{<}} Q(m) + \sum_{m \in M_{>}} Q(m) \epsilon_{\text{ext}} \leq \sum_{m \in M_{<}} \frac{\epsilon'}{m_{\max}} + \sum_{m \in M_{>}} Q(m) \epsilon_{\text{ext}} \leq \epsilon' + \epsilon_{\text{ext}}. \quad (35)$$

Since the actual distribution  $P_{\mathbf{X}\mathbf{V}E}$  characterizing the output of the device is  $\epsilon$ -close to  $Q_{\mathbf{X}\mathbf{V}E}$ , it directly follows that it provides an  $(\epsilon + \epsilon' + \epsilon_{\text{ext}})$ -secure realization of the protocol. Indeed, by the triangle inequality, and the fact that classical processing can only reduce the trace distance, we find  $d(P_{RFM}, \tilde{Q}_{RFM}) \leq d(P_{RFM}, Q_{RFM}) + d(Q_{RFM}, \tilde{Q}_{RFM}) \leq \epsilon + \epsilon' + \epsilon_{\text{ext}}$ . By the same argument, the protocol is  $(\epsilon_{\text{inp}} + \epsilon + \epsilon' + \epsilon_{\text{ext}})$ -secure when errors inherent to the input generation are taken into account (see Appendix A for an analysis of the errors introduced at this stage).

More generally, the security (in the context of classical side-information) of more complex protocols, such as those considered in [25, 26], where outputs of one pair of devices are used as inputs for another pair of devices can directly be proven from Theorem 1 and by keeping track of the error propagation.

## Efficiency

While the protocol presented above produces new randomness, it also uses a source of initial randomness  $S = (S_{\text{inp}}, S_{\text{ext}})$  to generate the inputs  $\mathbf{V}$  and perform the final randomness extraction. As a straightforward generalization of condition (10), the security of the protocol requires

this initial seed to be uniform and independent from the initial state of the devices, i.e.,

$$\rho_{SABE} = \omega_S \otimes \rho_{ABE} \quad (36)$$

where  $\omega_S$  denotes the uniform distribution on  $S$ .

This condition is obviously satisfied if  $S$  represents the output of a genuine, cryptographically-secure random number generator. Of course, a device-independent randomness expansion protocol is useful only if it produces more randomness at its output than is consumed at its input. It is shown in [14] how the protocol that we have presented above can achieve quadratic expansion by choosing appropriately the probabilities  $p_{vw}$  characterizing the input distribution. It can also be used as a primitive in more elaborate protocols where the output of one pair of devices are repeatedly used as input for another pair of devices. Such protocols can achieve exponential expansion, see [25] and particularly Section 5 of [26] for quantitative details (note that the application of our results, valid against classical-side information, to such concatenated protocols require not only that different pairs of devices be unentangled to the adversary to start with, but also between themselves. This assumption is again very reasonable in a trusted-provider situation).

Note, however, that to generate private randomness, a device-independent protocol does not necessarily need to consume any cryptographically-secure randomness to start with. Indeed, since we assumed in the security analysis that  $S$  was made public, the seed  $S$  does not need to be random with respect to the adversary, provided that condition (36) is satisfied, i.e., provided that the adversary cannot exploit any prior knowledge about  $S$  to influence the behaviour of the devices. If this is the case, which may be reasonable to assume in a trusted provider situation<sup>6</sup>, the output of the protocol will nevertheless represent randomness that is private with respect to the adversary.

## Acknowledgements.

We thank Ll. Masanes for pointing out an error in a previous version of this paper and Serge Fehr and Christian Schaffner for useful discussions. This work was supported by the European EU QCS project, the CHIST-ERA DIQIP project, the Interuniversity Attraction Poles Photonics@be Programme (Belgian Science Policy), the Brussels-Capital Region through a BB2B Grant.

## Appendix

Here we prove that one can use a uniform distribution to efficiently sample with exponentially small error from an i.i.d. non-uniform distribution, see also [45, 46].

**Theorem 2.** *Consider the finite alphabet  $K = a_1, \dots, a_{|K|}$ . Let  $Q$  be a probability distribution on  $K$  with  $\min_a Q(a) = n^{-\gamma}$ . Let  $a^n = a_1, a_2, \dots, a_n \in K^n$  be drawn i.i.d. according to  $Q$ . We denote  $Q^n$  the corresponding probability distribution on  $K^n$ . Suppose that  $x \in \{0, 1\}^m$  is drawn from the uniform distribution  $\omega$  on  $m$  bits. Then, for any  $0 \leq \gamma < 1/3$ , one can construct a function  $f : \{0, 1\}^m \rightarrow K^n$  such that the induced probability distribution on  $K^n$  given by  $P(a^n) = \omega(f^{-1}(a^n))$  is  $\epsilon$  close to  $Q^n$ , i.e.,  $d(P, Q^n) = \frac{1}{2} \sum_{a^n} |P(a^n) - Q^n(a^n)| \leq \epsilon$  with  $m \geq nH(Q) + o(nH(Q))$  and  $\epsilon \leq 3 \exp[-2n^{1-3\gamma}]$ , where  $H(Q) = \sum_a Q(a) \ln Q(a)$  is the Shannon entropy of  $Q$ .*

---

<sup>6</sup>Note though that even in a trusted provider situation, the condition (36) may fail, if the adversary can modify the behaviour of the devices by controlling external parameters like, e.g., the power supply of the devices.



*Proof.* The proof follows from Lemmas 4, 5, 6 below. Lemma 4 shows that there is a probable subset of  $K^n$  which occurs with high probability, and Lemma 5 computes the size of this probable subset. In Lemmas 4 and 5, we take parameter  $\alpha = n^{1/2-\gamma}$ . With this choice, from Lemma 4, the error one makes is  $\leq 2 \exp[-2n^{1-3\gamma}]$  and from Lemma 5 the size of the probable subset is  $\leq 2^{nH(Q)+O(n^{1-\gamma})}$ . Finally Lemma 6 tells us how one can sample efficiently from a distribution of known size. We take the error parameter in Lemma 6 to be  $\exp[-2n^{1-3\gamma}]$  (i.e. the same as in Lemma 4). The additional size penalty is negligible compared to the one coming from Lemma 5. This proves the result.  $\square$

*Counting Typical sequences.* Consider the alphabet  $K = a_1, \dots, a_{|H|}$ . If  $a^n = a_1, a_2, \dots, a_n \in K^n$  is a word of length  $n$  we denote by  $N(a|x)$  = number of occurrences of  $a \in K$  in word  $a^n$  (this is known as the type of the sequence). Let  $Q$  be a probability distribution on  $K$ . Let  $a^n = a_1, a_2, \dots, a_n \in K^n$  be drawn i.i.d. according to  $Q$ . We denote  $Q^n$  the corresponding probability distribution on  $K^n$ .

For any  $\alpha > 0$  define the set:

$$T_{Q\alpha}^n = \left\{ x \in K^n : \forall a \in K \ |N(a|x) - nQ(a)| \leq \alpha\sqrt{n}\sqrt{Q(a)} \right\}$$

**Lemma 4.**  $Q^n(T_{Q\alpha}^n) \geq 1 - 2|K| \exp[-2\alpha^2 \min_a Q(a)]$ .

*Proof.*  $T_{Q\alpha}^n$  is the intersection of  $|K|$  events, namely that for each  $a \in K$  the mean of the i.i.d Bernoulli variables  $y_i$ , defined by  $y_i = 1$  iff.  $a_i = a$  and  $y_i = 0$  iff.  $a_i \neq a$ , has deviation from its expected value  $Q(a)$  by at most  $\alpha\sqrt{n}\sqrt{Q(a)}$ . By the Hoeffding bound, each of these events has probability  $\geq 1 - 2 \exp[-2\alpha^2 Q(a)]$ . Hence the intersection of the events has probability  $\geq 1 - 2|K| \exp[-2\alpha^2 \min_a Q(a)]$ .  $\square$

**Lemma 5.**  $|T_{Q\alpha}^n| \leq 2^{nH(Q)+2\frac{\log_2 e}{\alpha}|K|\alpha\sqrt{n}}$ .

*Proof.* Consider  $x \in T_{Q\alpha}^n$ . Then  $Q(x) = \prod_{a \in K} Q(a)^{N(a|x)}$ . Hence

$$\begin{aligned} |-\log_2 Q(x) - nH(Q)| &= \left| \sum_{a \in K} -N(a|x) \log_2 Q(a) - nH(Q) \right| \\ &\leq \sum_{a \in K} -\log_2 Q(a) |N(a|x) - nQ(a)| \\ &\leq \sum_{a \in K} -\log_2 Q(a) \alpha\sqrt{Q(a)}\sqrt{n} \\ &= 2\alpha\sqrt{n} \sum_{a \in K} -\log_2 \sqrt{Q(a)} \sqrt{Q(a)} \\ &\leq 2\alpha\sqrt{n} \frac{\log_2 e}{e} |K| \end{aligned}$$

Therefore  $Q(x) \geq 2^{-nH(Q)-2\frac{\log_2 e}{\alpha}|K|\alpha\sqrt{n}}$ , and  $1 \geq \sum_{x \in T_{Q\alpha}^n} Q(x) \geq |T_{Q\alpha}^n| 2^{-nH(Q)-2\frac{\log_2 e}{\alpha}|K|\alpha\sqrt{n}}$  which proves the result.  $\square$

*Sampling from arbitrary distributions.* Suppose that  $x \in \{0, 1\}^m$  is drawn from the uniform distribution  $\omega$ .

Consider the probability distribution  $P(z)$  on  $z \in \{0, 1\}^k$ . We want to use  $x$  to sample with high precision from  $P(z)$ . That is, we define a function  $f : \{0, 1\}^m \rightarrow \{0, 1\}^k : x \rightarrow f(x)$  such that the induced probability distribution  $P'(z) = \omega(f^{-1}(x))$  is close to  $P(z)$ , as measured by the trace distance  $d(P, P') = \frac{1}{2} \sum_z |P(z) - P'(z)|$ . We have:

**Lemma 6.** For any  $\epsilon > 0$ , if  $m \geq k + \log_2 \frac{1}{\epsilon}$  we can construct a function  $f$  such that  $d(P, P') \leq \epsilon$ .

*Proof.* We view any  $x \in \{0, 1\}^m$  as a number in  $[0, 1]$  written in binary:  $x = \sum_{i=1}^m x_i 2^{-i}$ .

We define  $P'(z) \in \{0, 1\}^m$  as the largest binary number smaller than  $P(z)$ . Therefore  $0 \leq P(z) - P'(z) \leq 2^{-m}$ . We have  $1 - \sum_z P'(z) = \sum_z P(z) - P'(z) \leq 2^{-(m-k)}$ . To have a normalised distribution we define an additional outcome  $\perp$  with  $P'(\perp) = 1 - \sum_z P'(z)$ . Using  $x \in \{0, 1\}^m$  drawn from the uniform distribution  $\omega$ , we can therefore sample from  $P'(z)$  thus defined with  $d(P, P') = \frac{1}{2} \sum_z |P(z) - P'(z)| + \frac{1}{2} P'(\perp) \leq 2^{-(n-k)}$ . (The function  $f$  can be explicitly defined through  $f^{-1}(z) = \left\{ x : \sum_{z'=0}^z P'(z') \leq x \leq \sum_{z'=0}^{z+2^{-k}} P'(z') \right\}$ ).  $\square$

## References

- [1] T. Jennewein, U. Achleitner, G. Weihs, H. Weinfurter, and A. Zeilinger, A fast and compact quantum random number generator, *Review of Scientific Instruments* **71**, 1675 (2000).
- [2] A. Stefanov, N. Gisin, O. Guinnard, L. Guinnard, and H. Zbinden, Optical quantum random number generator, *Journal of Modern Optics* **47**, 595 (2000).
- [3] J. F. Dynes, Z. L. Yuan, A. W. Sharpe, and A. J. Shields, A high speed, postprocessing free, quantum random number generator, *Applied Physics Letters* **93**, 031109 (2008).
- [4] M. Fiorentino, C. Santori, S. M. Spillane, R. G. Beausoleil, and W. J. Munro, Secure self-calibrating quantum random-bit generator, *Physical Review A* **75**, 032334 (2007).
- [5] B. Qi, Y.-M. Chi, H.-K. Lo, and L. Qian, *High-speed quantum random number generation by measuring phase noise of a single-mode laser*, *Optics Letters* **35**, 312 (2010).
- [6] M. Fürst, H. Weier, S. Nauerth, D. G. Marangon, C. Kurtsiefer, and H. Weinfurter, *High speed optical quantum random number generation*, *Optics Express* **12**, p. 13029 (2010).
- [7] M. A. Wayne and P. G. Kwiat, *Low-bias high-speed quantum random number generator via shaped optical pulses*, *Optics Express* **18**, 9351 (2010).
- [8] Wikipedia, *Random number generation — Wikipedia, The Free Encyclopedia*, [http://en.wikipedia.org/w/index.php?title=Random\\_number\\_generation&oldid=470864939](http://en.wikipedia.org/w/index.php?title=Random_number_generation&oldid=470864939), accessed January 30, 2012.
- [9] Bell, J. S. *Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy* (Cambridge Univ. Press, 2004).
- [10] A. Valentini, *Signal-locality in hidden-variables theories*, *Physics Letters A* **297**, 273 (2002).
- [11] J. Barrett, L. Hardy, and A. Kent, No signaling and quantum key distribution, *Physical Review Letters* **95**, 010503 (2005).
- [12] Ll. Masanes, A. Acín, N. Gisin, *General properties of nonsignaling theories*, *Physical Review A* **73**, 012112 (2006).
- [13] R. Colbeck, *Quantum and Relativistic Protocols for Secure Multi-Party Computation*. PhD dissertation, Univ. Cambridge (2007), arXiv:0911.3814.
- [14] S. Pironio et al, *Random Numbers Certified by Bell's Theorem*, *Nature* **464**, 1021 (2010).

- [15] A. Ekert, *Quantum cryptography based on Bells theorem*, Physical Review Letters **67**, 661 (1991).
- [16] D. Mayers and A. Yao, *Self testing quantum apparatus*, Quantum Information and Computation **4**, 273 (2004).
- [17] A. Acín, N. Gisin, and Ll. Masanes, *From Bell's Theorem to Secure Quantum Key Distribution*, Physical Review Letters **97**, 120405 (2006).
- [18] A. Acín *et al.*, *Device-Independent Security of Quantum Cryptography against Collective Attacks* Physical Review Letters **98**, 230501 (2007).
- [19] J. Silman *et al.*, *Fully Distrustful Quantum Bit Commitment and Coin Flipping*, Physical Review Letters **106**, 220501 (2011).
- [20] C.-E. Bardyn *et al.*, Phys. Rev. A **80**, 062327 (2009).
- [21] J.-D. Bancal, N. Gisin, Y.-C. Liang, and S. Pironio, Phys. Rev. Lett. **106** 250404 (2011)
- [22] F. Magniez *et al.*, in *Proceedings of the 33rd International Colloquium on Automata, Languages and Programming*, (Springer, 2006), p. 72.
- [23] S. Fehr, R. Gelles, C. Schaffner, *Security and Composability of Randomness Expansion from Bell Inequalities*, arXiv:1111.6052v3.
- [24] U. Vazirani, T. Vidick, *Certifiable Quantum Dice Or, testable exponential randomness expansion*, arXiv:1111.6054
- [25] S. Pironio and S. Massar, arXiv:1111.6056v1.
- [26] S. Fehr, R. Gelles, C. Schaffner, *Security and Composability of Randomness Expansion from Bell Inequalities*, arXiv:1111.6052v2.
- [27] S. Pironio, *Device-Independent Quantum Key Distribution and Randomness Generation*, talk given at “Theory and Realization of Practical Quantum Key Distribution 2010”, Waterloo, June 16, 2010, slides available at [http://homepages.ulb.ac.be/~spironio/talks/Waterloo\\_2010.pdf](http://homepages.ulb.ac.be/~spironio/talks/Waterloo_2010.pdf).
- [28] V. Scarani and C. Kurtsiefer, *The black paper of quantum cryptography: real implementation problems*, arXiv:0906.4547.
- [29] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov, *Hacking commercial quantum cryptography systems by tailored bright illumination*, Nature Photonics **4**, 686 (2010).
- [30] Y. Zhao, C.-H. F. Fung, B. Qi, C. Chen, and H.-K. Lo, *Quantum hacking: Experimental demonstration of time-shift attack against practical quantum-key-distribution systems* Phys. Rev. A **78**, 042333 (2008).
- [31] F. Xu, B. Qi, and H.-K. Lo *Experimental demonstration of phase-remapping attack in a practical quantum key distribution system*, New J. Phys. **12**, 113026 (2010).
- [32] J. Barrett, R. Colbeck, A. Kent, *Prisoners of their own device: Trojan attacks on device-independent quantum cryptography*, arXiv:1201.4407v3.

- [33] *HotBits: Genuine random numbers, generated by radioactive decay*, available at [www.fourmilab.ch/hotbits](http://www.fourmilab.ch/hotbits).
- [34] *Random.org*, available at [www.random.org](http://www.random.org).
- [35] *EntropyPool and Entropy Filter Home Page*, available at [random.hd.org/index.html](http://random.hd.org/index.html).
- [36] R. Shaltiel, *An Introduction to Randomness Extractors, Automata, Languages and Programming*, Lecture Notes in Computer Science, volume 6756, p.21 (2011).
- [37] A. De, C. Portmann, T. Vidick, and R. Renner, *Trevisan's extractor in the presence of quantum side information*, arXiv:0912.5514.
- [38] R. Koenig, B. Terhal, *The Bounded Storage Model in The Presence of a Quantum Adversary*, IEEE Trans. Inf. Th. **54**, no. 2 (2008).
- [39] A. Acin, S. Massar, S. Pironio, *Randomness vs non-locality and entanglement*, arXiv:1107.2754
- [40] M. Navascués, S. Pironio, A. Acín, Phys. Rev. Lett. **98**, 010401 (2007); New J. Phys. **10**, 073013 (2008).
- [41] S. Pironio, M. Navascués, A. Acín, SIAM J. Optim. vol **20**, issue 5, 2157 (2010).
- [42] G. Grimmett and D. Stirzaker, *Probability and Random Processes*, Oxford University Press, 2001.
- [43] M. Ben-Or, M. Horodecki, D. W. Leung, D. Mayers, and J. Oppenheim, *The universal composable security of quantum key distribution*, Lecture Notes in Computer Science, Volume 3378, 386-406 (2005)
- [44] M. Tomamichel, C. C. W. Lim, N. Gisin, R. Renner, *Tight finite-key analysis for quantum cryptography*, arXiv:1103.4130 (2011).
- [45] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, J. Wiley (1991), chapter 5.12.
- [46] D. Knuth and A. Yao, *Algorithms and Complexity: New Directions and Recent Results*, chapter *The complexity of nonuniform random number generation*, Academic Press, 1976.
- [47] R. Colbeck and A. Kent, *Private Randomness Expansion With Untrusted Devices*, Journal of Physics A: Mathematical and Theoretical **44**, 095305 (2011).